

Speech analysis and synthesis systems for the tatar language

Khusainov A., Khusainova A.

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2016 FRUCT. In this paper we describe our recent work of creation speech human-machine interface for the Tatar language. Our work consists of three main elements: speech recognition system, speech synthesizer and language identification system. These systems will be used in mobile and desktop applications, for instance, machine translation system, smart assistant.

References

- [1] A. Khusainov, "Design and creation of speech corpora for the Tatar speech recognition and synthesis tasks", Proc. of the 3rd International Conference on Turkic Languages Processing, Kazan, 2015, pp. 475-484.
- [2] VoxForge official website, Web: <http://www.voxforge.org/>.
- [3] LDC official website, Language resources, Web: <https://catalog.ldc.upenn.edu/Ldc93sl>.
- [4] HTK speech recognition toolkit, Web: <http://htk.eng.cam.ac.uk/>.
- [5] A.F. Khusainov, "Automatic phoneme recognition system for the Tatar language", Proc. Of the 1st International Conference on Turkic Languages Processing, Astana, 2013, pp 211-217.
- [6] A. Stolcke, "SRILM-An Extensible Language Modeling Toolkit", Proc. Intl. Conf. on Spoken Language Processing, vol. 2, Denver, 2002, pp. 901-904.
- [7] D. Sh. Suleymanov, R. A. Guilmouline, A. A. Guilmouline, "Tatar phonological rules as a base of two-level morphological analyzer", in Proceedings of LP'2000, Prague: The Karolinum Press.-P. 495504.
- [8] M. Creutz, K. Lagus, "Unsupervised discovery of morphemes", In Proceedings of the Workshop on Morphological and Phonological Learning of ACL-02, Philadelphia, Pennsylvania, 11 July, 2002, pp. 21-30.
- [9] Dz. Suleymanov, O. A. Nevzorova, and B. Khakimov, "National Corpus of the Tatar Language "Tugan Tel": Structure and Features of Grammatical Annotation", Proc. International Conference Georgian Language and modern Technology, Tbilisi, 2013, pp. 107-108.
- [10] P. F. Brown, V. J. Delia Pietra, P. V. deSouza, J. C. Lai and R. L. Mercer, "Class-Based n-gram Models of Natural Language", Computational Linguistics, vol. 18 (4), 1992, pp. 467-479.
- [11] Open-Source Large Vocabulary Continuous Speech Recognition Engine, Web: <https://github.com/julius-speech/julius>.
- [12] Programmye produkty, lokalizovannye na tatarskiy yazyk, Web: <http://tatsoft.tatar>.
- [13] E. Moulines, F. Charpentier, "Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones", In Speech Communication, 9 (5/6), 1990, pp. 453-167.
- [14] Y. Sagisaka, "ATR v-talk speech synthesis system", In Proc. ICSLP-92, Banff, Canada, 1992.
- [15] T. Yoshimura, K. Tokuda, T. Masuko, T. Kobayashi, T. Kitamura, "Simultaneous modeling of spectrum, pitch and duration in HMM-based speech synthesis", In Proc. Eurospeech, 1999, pp. 2347-2350.